Estimation of Heart Rate from Photoplethysmography during Physical Exercise using Wiener Filtering and the Phase Vocoder

Andriy Temko, Senior Member, IEEE

Abstract—A system for estimation of the heart rate (HR) from the photoplethysmographic (PPG) signal during intensive physical exercises is presented. The Wiener filter is used to attenuate the noise introduced by the motion artifacts in the PPG signals. The frequency with the highest magnitude estimated using Fourier transformation is selected from the resultant de-noised signal. The phase vocoder technique is exploited to refine the frequency estimate, from which the HR in beats per minute (BPM) is finally calculated. On a publicly available database of twenty three PPG recordings, the proposed technique obtains an error of 2.28 BPM. A relative error rate reduction of 18% is obtained when comparing with the state-of-the-art PPG-based HR estimation methods. The proposed system is shown to be robust to strong motion artifact, produces high accuracy results and has very few free parameters, in contrast to other available approaches. The algorithm has low computational cost and can be used for fitness tracking and health monitoring in wearable devices.

I. INTRODUCTION

The functionality of wearable devices has been gradually increasing over the last decades. Accurate heart rate (HR) monitoring during physical exercise is one of many useful features that modern wearable devices can offer. Implemented in smart-watches or wristbands, it can guide exercisers to adapt their training load and better match their training goals. The task of losing weight requires maintaining a heart-rate which is different to that of aerobic exercises. The current fitness equipment can dynamically adjust the parameters, in contrast to other available approaches. The algorithm has low computational cost and can be used for fitness tracking and health monitoring in wearable devices.

One of the most unobtrusive ways to estimate HR in real-time is to use photoplethysmographic (PPG) signals. These signals are usually recorded from the wrist. The technology of the PPG signals [1-3] relies on data from pulse oximeters which are embedded in these wearable devices. A pulse oximeter emits light to the skin and measures the changes of intensity of the light which is reflected from the skin. These measures sense the rate of blood flow as controlled by the heart’s pumping action. The periodicity of these measurements in most cases corresponds to the cardiac rhythm, and thus, HR can be estimated from the PPG signal. However, motion artifacts (MA) are known to be a limiting factor that prevents the straight-forward usage of PPG in free living conditions. Strong movements during physical exercise makes the HR estimate inaccurate (Fig. 1). A number of methods have been proposed to remove or attenuate MA in PPG signals, including adaptive filtering [4, 5], independent component analysis [6], empirical mode decomposition [7], spectral subtraction [8], Kalman filtering [9], to name a few. Some of these routines exploit the existence of accelerometer data whilst others are tested under mild-movement conditions only where MA are not strong (walking, finger movements).

The TROIKA method has recently been proposed to estimate HR from PPG signals for scenarios where MAs are strong [2]. The method was based on signal decomposition, sparsity-based high-resolution spectrum estimation, and spectral peak tracking and verification. The average absolute error of 2.34 beats per minute (BPM) was reported on 12 PPG recordings. The TROIKA method was modified in [3]. The spectra of PPG and acceleration signals were jointly estimated using a common sparsity constraint on the spectral coefficients by means of a multiple measurement vector model [10]. The error was reduced to 1.28 BPM when evaluated on the same 12 PPG recordings.

It is difficult to verify genuine technical contributions when the databases, metrics reported, and performance assessment routines used vary from study to study. However, for the IEEE Signal Processing Cup 2015 the database of 23 PPG recordings was made public. The challenge also provided the definition of the rules that can be followed and the metrics that should be computed to compare different approaches.

Figure 1. The challenge of HR estimation during physical exercises. Plot (a) and (b) show the clean PPG waveform and its periodogram. Plot (c) and (d) show the waveform and the periodogram of the PPG which is corrupted with motion artifacts. The true heart rate is denoted with red circle. Taken from [14].
from 18 to 58 years old subjects performing various physical exercises. For each subject, the PPG signals were recorded and sent to a nearby computer via Bluetooth. The ECG signal was recorded simultaneously from the chest using wet ECG sensors. All signals were sampled at 125 Hz (wavelength: 515nm). The acceleration signal was also recorded from the wrist using a three-axis accelerometer. Both the pulse oximeter and the accelerometer were embedded in a wristband, which was comfortably worn. The developed system is described in detail in Section 3. The results on the provided data are presented in Section 4.

### TABLE I. DATABASE OF PPG RECORDINGS.

<table>
<thead>
<tr>
<th>Rec</th>
<th>Subject ID</th>
<th>Activity type</th>
<th>Age/Weight/Height</th>
<th>Sex</th>
<th>Healthy?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>T1</td>
<td>18-35y/-/-</td>
<td>M</td>
<td>Y</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>T1</td>
<td>20y/64kg/162cm</td>
<td>M</td>
<td>Y</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>T1</td>
<td>21y/77kg/188cm</td>
<td>M</td>
<td>Y</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>T1</td>
<td>22y/77kg/188cm</td>
<td>M</td>
<td>Y</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>T1</td>
<td>23y/54kg/174cm</td>
<td>M</td>
<td>Y</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>T1</td>
<td>24y/54kg/174cm</td>
<td>M</td>
<td>Y</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>T1</td>
<td>25y/57kg/174cm</td>
<td>M</td>
<td>Y</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>T1</td>
<td>26y/70kg/180cm</td>
<td>M</td>
<td>Y</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>T1</td>
<td>27y/70kg/180cm</td>
<td>M</td>
<td>Y</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>T1</td>
<td>28y/70kg/180cm</td>
<td>M</td>
<td>Y</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>T1</td>
<td>29y/70kg/180cm</td>
<td>M</td>
<td>Y</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>T1</td>
<td>30y/64kg/162cm</td>
<td>M</td>
<td>Y</td>
</tr>
<tr>
<td>13</td>
<td>13</td>
<td>T2</td>
<td>31y/64kg/162cm</td>
<td>M</td>
<td>Y</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>T2</td>
<td>32y/70kg/180cm</td>
<td>M</td>
<td>Y</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>T2</td>
<td>33y/70kg/180cm</td>
<td>M</td>
<td>Y</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>T3</td>
<td>34y/54kg/174cm</td>
<td>M</td>
<td>Y</td>
</tr>
<tr>
<td>17</td>
<td>17</td>
<td>T3</td>
<td>35y/54kg/174cm</td>
<td>M</td>
<td>Y</td>
</tr>
<tr>
<td>18</td>
<td>18</td>
<td>T3</td>
<td>36y/57kg/174cm</td>
<td>M</td>
<td>Y</td>
</tr>
<tr>
<td>19</td>
<td>19</td>
<td>T3</td>
<td>37y/70kg/180cm</td>
<td>M</td>
<td>Y</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td>T2</td>
<td>38y/70kg/180cm</td>
<td>M</td>
<td>Y</td>
</tr>
<tr>
<td>21</td>
<td>21</td>
<td>T3</td>
<td>39y/70kg/180cm</td>
<td>M</td>
<td>Y</td>
</tr>
<tr>
<td>22</td>
<td>22</td>
<td>T3</td>
<td>40y/70kg/180cm</td>
<td>M</td>
<td>Y</td>
</tr>
<tr>
<td>23</td>
<td>23</td>
<td>T2</td>
<td>41y/70kg/156cm</td>
<td>F</td>
<td>HR&amp;BP</td>
</tr>
</tbody>
</table>

T1 = walking/running on a treadmill. T2 = rehabilitation arm exercises. T3 = intensive arm movements (boxing)

In this paper a new approach to HR estimation is proposed, which is based on Wiener filtering and the phase vocoder. The noise signature is estimated from accelerometer signals and the Wiener filter is used to attenuate the noise components in the PPG signal. Additionally, the phase vocoder is exploited to overcome the limited resolution of the Fourier transform and to refine the initial dominant frequency estimation. In contrast to the previously presented system, the proposed solution requires only a few parameters to tune.

The paper is organized as follows. Section 2 describes the database, metrics and the performance assessment routine used in the study. The developed system is described in detail in Section 3. The results on the provided data are presented in Section 4.

### II. DATA AND TASK SPECIFICATIONS

#### A. Database

The data used in this work were provided for the IEEE Signal Processing Cup and are publically available. The dataset consists of 23 5-min recordings which were collected from 18 to 58 years old subjects performing various physical exercises. For each subject, the PPG signals were recorded from the wrist using two pulse oximeters with green LEDs (wavelength: 515nm). The acceleration signal was also recorded from the wrist using a three-axis accelerometer. Both the pulse oximeter and the accelerometer were embedded in a wristband, which was comfortably worn. The ECG signal was recorded simultaneously from the chest using wet ECG sensors. All signals were sampled at 125 Hz and sent to a nearby computer via Bluetooth.

The details of the database are given in Table I. Three types of activities were performed. Type 1 (T1) activity involved walking or running on a treadmill for intervals of 0.5m-1m-1m-1m-0.5m with the speed of 1-2 km/h, 6-8 km/h, 12-15 km/h, 6-8 km/h, 12-15 km/h, 1-2 km/h, respectively. The subjects were asked to purposely use the hand with the wristband to pull clothes, wipe sweat on forehead, and push buttons on the treadmill. Type 2 activity included various forearm and upper arm exercise which are common in arm rehabilitation (e.g. shake hands, stretch, push, running, jump, and push-ups). Type 3 activity consisted of intensive forearm and upper arm movements (e.g. boxing).

The ECG-based ground truth HR was provided to assess the performance of the developed algorithms.

The first 12 of the 23 recordings were used in the evaluation of the TROIKA [2] and JOSS [3] PPG-based HR estimation systems. This allows the direct comparison with the results in this study. The other 11 were released specifically for the IEEE Signal Processing Cup 2015.

#### B. Task specifications

Several rules were defined for the competition and are followed in this study. The developed algorithms are allowed to use any number of signals provided in the data. For example, the algorithms may choose to use a single PPG signal, or use only PPG signals without accelerometer data for HR estimation.

The HR must be estimated from an 8s-window every 2s. The algorithms are allowed to use the previous data up to the current point in time such as previous HR estimates or past samples. However, it is not allowed to use any data from the future so that no algorithmic delay in the processing chain is allowed. For instance, this prohibits the usage of any non-causal smoothing operators such as central moving average filters or modification of the past HR estimates based on the current HR estimate, etc.

#### B. Metrics

The Average Absolute Error (AAE) is used to evaluate the performance. It is defined as:

$$ AAE = \frac{1}{N} \sum_{i=1}^{N} \left| BPM_{\text{est}}(i) - BPM_{\text{true}}(i) \right| $$

where $N$ is the total number of estimates (number of windows), $BPM_{\text{est}}(i)$ and $BPM_{\text{true}}(i)$ denote the estimated and the true HR value in the $i$-th time window in BPM, respectively. Similarly, the Standard Deviation of the Absolute Error is computed (SDAE). The average of these metrics across all 23 recordings is reported.

### III. METHOD

The flowchart of the developed system is shown in Fig. 2. During preprocessing two PPG signals and 3 accelerometer signals are filtered with the 4th order Butterworth band-pass filter (0.4-4Hz). The two PPG signals are then normalized to zero mean and unit variance and averaged. The averaged PPG signal and 3 accelerometer signals are down-sampled from 125 to 25Hz for further processing.

The signals are then subjected to the Discrete Fourier Transform (DFT). The content that corresponds to the HR between 60 BPM and 180 BPM is kept. Wiener filtering is applied to remove the MA from the PPG signal. In the

1 http://zhilinzhang.com/spcup2015/data.html
frequency domain, the noisy PPG signal, $Y(f)$, is assumed to be corrupted by additive MA noise:

$$Y(f) = X(f) + N(f)$$  \hspace{1cm} (2)$$

where $X(f)$ and $N(f)$ are the spectra of the clean PPG signal and the MA which are captured by the accelerometer signals. For a signal observed in uncorrelated additive random noise, the frequency-domain Wiener filter is given as:

$$W(f) = \frac{P_{XX}(f)}{P_{XX}(f) + P_{NN}(f)}$$  \hspace{1cm} (3)$$

where $P_{XX}(f)$ and $P_{NN}(f)$ are the signal and noise power spectra. Thus, the Wiener filter acts as a signal-to-noise dependent attenuator. The filter requires separate estimates of the noise and signal power spectrums. The noise spectrum can be estimated from accelerometer signals whereas the clean PPG spectrum, $P_{XX}(f)$, can be estimated as $P_{Y}(f)-P_{NN}(f)$ or recursively from previous filter outputs. Depending on the manner the power spectrum of the clean PPG signal is estimated two Wiener filters are implemented:

$$w_1(t,k) = 1 - \frac{P_{NN}(t,k)}{C \sum_{i=-C+1}^{t} P_{YY}(i,k)}$$  \hspace{1cm} (4)$$

$$w_2(t,k) = \frac{1}{C \sum_{i=-C+1}^{t} P_{YY}(i,k)w_2(i,k)} - \frac{P_{NN}(t,k)}{C \sum_{i=-C+1}^{t} P_{YY}(i,k)w_2(i,k) + P_{NN}(t,k)}$$  \hspace{1cm} (5)$$

where $w(t,k)$ is the weight of the $k$-th frequency bin at time, $t$. The noise is estimated instantaneously by averaging the spectrum from 3 accelerometer signals. After that the spectral envelopes are divided by their maximum value to assure commensurability between PPG and accelerometer signals. In Eq. 4, the power spectrum of the clean PPG signal is estimated by subtracting the observed noise from the observed PPG signal. The power spectrum of the observed signal is averaged over the past $C$ spectral envelopes ($C=15$, in this work). If $C=1$, then the Wiener filter performs simple spectral subtraction [16, 17]. In Eq. 5, the spectra of the clean PPG signal is computed recursively by averaging the previous filtered signal outputs.

The spectral envelopes processed with the two designed filters are scaled by their std because unlike $w_2$, $w_1$ can have negative values. The resultant signals are averaged to give a final spectral envelope of the cleaned PPG signal. The dominant frequency (the frequency with the highest magnitude) is converted to the HR estimate in BPM.

The effective frequency resolution of the DFT is limited by the ratio between the sampling rate (25Hz) and the window size (8s) which equals to 7.5 BPM. In order to increase the resolution of the DFT the phase vocoder technique [9–11] is used to estimate the instantaneous frequency. The phases from the chosen peak in the magnitude spectrum from the current and previous frames are used to refine the initial frequency estimation

$$\phi_{new} = \left(\theta_2 - \theta_1 + 2\pi n\right)/(2\pi(t_2-t_1))$$  \hspace{1cm} (6)$$

where $n$ is an integer, $\theta_2$, $\theta_1$ are the two phases from the current and previous frames, respectively; $t_2$, $t_1$ are the time stamps of the two frames, here $t_2-t_1 \approx 2s$. The series of $\phi_{new}$ is computed for several $n$ using Eq. 6, and the value that is closest to the initial frequency estimation is chosen.

The post-processing steps include history tracking and smoothing. The previous estimation of the heart rate is used to guide the search range for the maximum DFT magnitude in the current frame. Additionally, the weighted average between the current estimate and its prediction with the linear regression is computed to smooth the series of HR estimates if the difference between the current and previous HR values is above 5 BPM. Finally, a small offset which depends on the derivative of the HR is added to account for consistently lower or higher heart rate estimations. The algorithm has been developed in Matlab and is available for download.

IV. RESULTS AND DISCUSSION

The results of the proposed HR estimation system are shown in Table 2. The first 12 of the 23 recordings of the dataset were used in [2] where the AAE of 2.34 with the $SDAE$ of 2.47 BPM were reported. Subsequently, on the same dataset the error was reduced to 1.28 with the increased $SDAE$ of 2.61 in [3]. Evaluating on the same 12 recordings, the developed system obtains the AAE of 1.05 with the $SDAE$ of 1.31 which corresponds to the AAE/SDAE reduction of 18%/50% relative, respectively. It can be seen that the error was significantly reduced for the worst performing recordings in the dataset.

The $AAE$ ($SDAE$) achieved across all 23 recordings was 2.28 (3.33) BPM. It can be seen from the obtained results that the error depends on the type of physical activity. Running on a treadmill results in the AAE of 1.05 BPM, whereas arm exercises and intensive arm exercises result in the error of 3.68 and 3.56 BPM, respectively.

Fig. 3 and Fig. 4 illustrate the output of the system for the two recordings with the lowest and the highest errors, recording 9 (treadmill) and 14 (arm exercise). It can be seen that the largest error is obtained when both the spectral signatures of the HR and MA change rapidly so that the true HR cannot be reliably estimated. Other spectral estimation methods may lead to more accurate results.

\[ http://github.com/andtem2000/PPG \]
The approaches reported in [2] and [3] rely on a number of heuristic rules and thresholds. For instance, the post-processing step alone in [3] requires a dozen parameters to be specified. Increasing the number of degrees of freedom of the designed system will lead to improved performance. However, it comes at the cost of the increased risk of poor generalisation on the unseen data especially if highly-tuned rules are introduced in the post-processing to correct the errors of the core algorithm. In contrast, the method presented in this study requires only a few parameters to be tuned. The number of frequency bins in the DFT was set to 1024. Likewise, $C=15$ was used in the Wiener filters. The range of search for the maximum in DFT magnitudes for the next frame was set to be $\pm 18$ BPM from the current HR estimate.

The proposed algorithm provides a simple but effective solution to the PPG-based HR estimation and can serve as a baseline performance for further studies.

**ACKNOWLEDGMENT**

The author would like to thank Caitlin Keane, Matthew O’Donovan, and Dr. Emanuel Popovici for their help in conducting the study.

**REFERENCES**


